



Practitioner's Docket No. MPI01-019P1RNM

PATENT

In re application of:	Curtis, Rory A.J.		
Application No.:	10/074,547	Group No.:	1647
Filed:	February 12, 2002	Examiner:	Kapust, Rachel B.
For:	25466, A HUMAN TRANSPORTER FAMILY MEMBER AND USES THEREFOR		

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

DECLARATION UNDER 37 C.F.R. § 1.131

Sir:

I, Rory A.J. Curtis, hereby declare and state:

1. I am the inventor of the subject matter described and claimed in the above-identified application.
2. I was in possession of the complete nucleic acid sequence of 25466 in the United States and had determined its identity as a monocarboxylate transporter before January 30, 2001.
3. Evidence is provided by the following:
 - a) Prior to January 30, 2001, I had completed the sequencing of full length human 25466. Exhibit A is a copy of an electronic printout of the nucleic acid sequence from Millennium's sequence database. The printout bears the automatically embedded date on which I entered the sequence into the database. On the original printout, this sequence bears a date prior to January 30, 2001. In accordance with accepted practice, the dates on the copy of the electronic printout have been masked (M.P.E.P. § 715.07).

The top sections of the printout bear reference information for accessing the 25466

CERTIFICATION UNDER 37 C.F.R. SECTIONS 1.8(a) and 1.10*

I hereby certify that, on the date shown below, this correspondence is being:

MAILING

- ☒ deposited with the United States Postal Service in an envelope addressed to the Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.
37 C.F.R. SECTION 1.8(a)


37 C.F.R. SECTION 1.10*

- ☒ with sufficient postage as first class mail.

☐ as "Express Mail Post Office to Addressee"
Mailing Label No.

TRANSMISSION

- ☐ transmitted by facsimile to the Patent and Trademark Office.


Signature
Sean Hunziker/Beverly Sotiropoulos

Date: July 8, 2004

(type or print name of person certifying)

***WARNING:** Each paper or fee filed by "Express Mail" must have the number of the "Express Mail" mailing label placed thereon prior to mailing. 37 C.F.R. section 1.10(b). "Since the filing of correspondence under section 1.10 without the Express Mail mailing label thereon is an oversight that can be avoided by the exercise of reasonable care, requests for waiver of this requirement will not be granted on petition." Notice of Oct. 24, 1996, 60 Fed. Reg. 56,439, at 56,442.

sequence in the Millennium database. The remainder of the printout bears the entire 25466 nucleotide sequence ("Fbh25466pat"), known in the application as SEQ ID NO:1.

b) Prior to January 30, 2001, I had determined that the 25466 nucleotide sequence of SEQ ID NO:1 encodes a monocarboxylate transporter based on the results of analyses of the nucleotide sequence and the encoded polypeptide. Exhibits B, C and D are evidence of this determination. Exhibit B is a copy of the first page of a BLAST analysis of SEQ ID NO:1. The top hit in this BLAST result is human monocarboxylate transporter 3 (Genbank Accession No. U81800). Exhibit C is a copy of several analyses of the 25466 polypeptide, with the sequence displayed on page 1. Page 4 of Exhibit C shows the alignment of a portion of the polypeptide with the MCT (monocarboxylate transporter) domain in Pfam. Exhibit D is a pairwise alignment of the 25466 polypeptide with human monocarboxylate transporter 4 (SwissProt Accession No. O15374). I performed my analyses and reached the conclusion that SEQ ID NO:1 encodes a monocarboxylate transporter prior to January 30, 2001.

The original printouts of Exhibits B and C bear automatically embedded dates on which I performed these analyses. In accordance with accepted practice, the dates on the copies of the electronic printouts have been masked (M.P.E.P. § 715.07).

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Rory Curtis

Rory A.J. Curtis

6/12/2004

Date

Sample Information

name: Fbh25466pat
type: Plain DNA Sequence
submitted on:
NC project: Bayer
species: Human
tissue: unspecified
internal id: 107698647
restrictions: none

Exhibit A to Accompany Declaration under
37 CFR §1.131 for US Appln. No. 10/074,547

Sequences Information

creation method: Loading
name: Load
created on:
created by: Rory Curtis
accession number: 107698648

> Fbh25466pat - Import - complete

```
1      GTCGACCCAC GCGTCCGCAA GAGTGTGCAT GTGAGGTGAC TGCATTTTTT TTCCCTGCCA
61     AACCAGAATT AGCCGGTATA GGAATGAACG AGCATGAAGA TTTGAAATTG CTCCGATTGG
121    AAGGAAGCCC AGGTTAGGTT TGGGCACCTC CAAACGCACC CGTTTTAAAG CCACCTGGAC
181    TGAGGCGTCG AGCTTTCAGC TCCACCAAAC GCTCACCTGG CCTGGCAGCG AGCGGCGGAA
241    GAGCCCGGGA GCCCCTCACA GAGCGCACCG AGCCGGGCGG AGAGCTGAGC CGCAGGCACC
301    CGCGTCTCCA GGATGATAGG CGACATTGCA ACAAATCTCT ACACCCAGCA GCTCAGGGGG
361    CTCCAAGCAG AGCAGCAAGT TCGAGGATCC GGGCGTGGAG CCGAGTGAGG CCGCAGCCCA
421    GCGGGCCTCG GCGGAAAAAT CTTGGAATAA GTATACCAGT CATGAAGATA TTGGGTATGA
481    TTTTGAAGAT GGCCCCAAAG AAAAAAAGAC ACTGAAGCCC CACCCAAACA TTGATGGCGG
541    ATGGGCTTGG ATGATGGTGC TCTCCTCTTT CTTTGTGCAC ATCCTCATCA TGGGCTCCCA
601    GATGGCCCTG GGTGTCTCTA ACGTGAATG GCTGGAAGAA TTCCACCAGA GCCGCGGCCT
661    GACCGCCTGG GTCAGCTCCC TCAGCATGGG CATCACCTTG ATAGTGGGCC CTTTCATCGG
721    CTTGTTTCATT AACACCTGTG GGTGCCGCCA GACTGCGATC ATTGGAGGGC TCGTCAACTC
781    CCTGGGCTGG GTGTTGAGTG CCTATGCTGC AAACGTGCAT TATCTCTTCA TTACTTTTGG
841    AGTCGCAGCT GGCCTGGGCA GCGGGATGGC CTACCTGCCA GCGGTGGTCA TGGTGGGCAG
901    GTATTTCCAG AAGAGACGCG CCCTCGCCCA GGGCCTCAGC ACCACGGGGA CCGGATTCGG
961    TACGTTCTTA ATGACTGTGC TGCTGAAGTA CCTGTGCGCA GAGTACGGCT GGAGGAATGC
1021   CATGTTGATC CAAGGTGCCG TTTCCCTAAA CCTGTGTGTT TGTGGGGCGC TCATGAGGCC
1081   CCTCTCTCCT GGTAAAAACC CAAACGACCC AGGAGAGAAA GATGTGCGTG GCCTGCCAGC
1141   GCACTCCACA GAATCTGTGA AGTCAACTGG ACAGCAGGGA AGAACAGAAG AGAAGGATGG
1201   TGGGCTCGGG AACGAGGAGA CCCTCTGCGA CCTGCAAGCC CAGGAGTGCC CCGATCAGGC
1261   CGGGCACAGG AAGAACATGT GTGCCCTCCG GATTCTGAAG ACTGTCAGCT GGCTACCAT
1321   GAGAGTCAGG AAGGGCTTCG AGGACTGGTA TTCGGGCTAC TTTGGGACAG CCTCTCTATT
1381   TACAAATCGA ATGTTTGTAG CCTTTATTTT CTGGGCTTTG TTTGCATACA GCAGCTTTGT
1441   CATCCCTTC ATTCACTCC CAGAAATCGT CAATTTGTAT AACTTATCGG AGCAAAACGA
1501   CGTTTTCCCT CTGACGTCAA TTATAGCAAT AGTTCACATC TTTGAAAAG TGATCCTGGG
```

1561	CGTCATAGCC	GACTTGCCTT	GCATTAGTGT	TTGGAATGTC	TTCCTGTTGG	CCAACTTCAC
1621	CCTTGTCCCTC	AGTATTTTTA	TTCTGCCGTT	GATGCACACG	TACGCTGGCC	TGGCGGTCAT
1681	CTGTGCGCTG	ATAGGGTTTT	CCAGTGGTTA	TTTCTCCCTA	ATGCCCCTAG	TGACTGAAGA
1741	CTTGGTTGGC	ATTGAACACC	TGGCCAATGC	CTACGGCATC	ATCATCTGTG	CTAATGGCAT
1801	CTCTGCATTG	CTGGGACCAC	CTTTTGCAGG	GTGGATCTAT	GACATCACGC	AAAAATATGA
1861	TTTTTCCTTC	TACATATGTG	GTTTGCTTTA	CATGATAGGA	ATACTCTTTT	TACTTATTCA
1921	GCCGTGCATT	CGAATTATAG	AACAATCCAG	AAGAAAATAC	ATGGATGGTG	CACATGTTTA
1981	GTATCATGTA	ATGTTCCGTG	TAGGTTTCAT	TGTAATACTC	ATGCCTACCT	CGCATGGTTG
2041	CTGTGAGGCA	CCTATGACAG	GACGTGGGAA	AGCATTTTGT	ACGGTAACTG	GCACTGTCAT
2101	TTGTAAATGC	CATTGTCACA	GCCTCATTTG	TAAGCAGCAC	TGCCTCTCTG	TTTGGGGAGA
2161	TGTAATGCTG	GAAGATCTTA	AGGACTACAT	ACATTCTAGA	GATGACAGTG	TTGTTCAAAG
2221	ACAGCCTAGT	AAGTAAATTG	TAGAAATGCC	CTTATAAAAA	CCATTCTCTT	GTCATCTACT
2281	GGGACTAGGG	TTTTAAATAC	AGCTTTTAAA	AACAAAAACA	GGGAATAAAA	GCTTTTCAAC
2341	TCAACCACTT	CTTTGTAAGA	CAAACTGAA	GTATCTGTGT	GCTTCCAGAA	AGCTTACAGA
2401	TAAATGGGTT	TCAAGCACAA	GAATATGACT	AGATTTTCTA	AATTAATTAT	TACAGGGAGC
2461	TATTGATCTA	CTAGCATCAA	ACAAAGGCAA	GCTCTAATTC	CACAGGTAAT	ACAATTTAGT
2521	GCAATTAAAG	AAACACGGCT	TGTATTTTTA	TGAGGGAATT	CTGCAGCTAG	GGATTGTGAC
2581	TCCTAAATCC	TCCTCTAAAA	GAAGGCACTT	GCCATTAATC	CTAATTCAGT	GCTATCCAGT
2641	TATAAATGGA	ATCTTGAGAC	AAAACCTTAA	CAAAGAAATA	ACAGTAATGA	TTTCCTTAGC
2701	AGAAGCCGTA	TTTGTACGCA	CAACATTAAA	TCAAGGGCTA	CAATTCAAGC	ACTTTTATTC
2761	GTATCATTGG	CCTCTTAGAT	GATATAAGCA	TGAGGTGGGG	CCTGTAATAT	TTTTTTCTGA
2821	GTTTCTTCTG	CCCAAAAAATA	TAATATAGAA	CTAATTGCTA	ACTGACAAAT	AAAGTTAATA
2881	GTTAAATCAT	CTCCAAGGAA	TGTTGCTAAT	CCAAAGTATA	ACACTATCAA	TTTGTGAGGA
2941	TAATAAATGG	AATGCCATTA	GTGTAGATGT	CTGTGCCACA	TCTGACACTG	GAGTAGTGAT
3001	AACAAATAGC	CCATCTCTAG	ACTCTCGTGT	TGTTATATAG	ACCATTCAAT	TGCCTGAGCG
3061	TGGCACAGTT	TTAAAAATAG	TTCTCTTGAT	TGATTTTCTA	CAGAAGATGA	CTGTGATCCA
3121	TGACATCTAA	TAATGCCCTT	TCTTTATCTG	AGATGTCTAT	TTTTCTAAGC	CAAACGTTTT
3181	TCAGACTGCA	GAATGTTCTT	CCCAGATCAT	TTGAAATTTC	TGGCTGCCTT	ACTTGTTTAC
3241	AGATAGTTTA	AGACTATTTA	AATTTCTACT	CACAATTTGA	TCATCACACA	CACACAAATC
3301	CTTGAATATC	ATTGCCAGTG	TCTTAGGTCA	AATTTACCTA	AAGTGAATAC	AGCCCATTCT
3361	CAATTATCCT	TCACAATTAG	ACGCAGGAAT	GCTACTAGGA	ATTGGAATCA	AACAATGCCA
3421	CCCCAAGCGT	AATTTTAGCC	AGCAGTTTCA	GTTATACTCA	ACCATGTCCT	TCTGAGCTGT
3481	TAACAAGTGA	TTCAATGGAC	AAGTTCTCTT	TTTGTTCCTT	CTCCATTATT	TCCTGCTCTA
3541	ATGTATAGTG	GGAGTGGTTG	TGTAATGAAA	GGACCACCAA	AATAATAAAA	GGCAGCTAAT
3601	GGAAAGGAGA	GACAAAAGCA	TGGTTAATAT	ATATACTTAA	TATTACCTCC	AATGACTCGG
3661	GAATTGCCTG	TAAATTATTA	TAGACAATAG	ATTGCATGTC	ATACTCCATT	TGGTTCAACA
3721	CAACAACCTA	TGTGTTATCA	TTACAGCTTT	GGCTGCTGTT	AAAGAATCCA	GCTCTCTATT
3781	TTGATAAAGA	TAATCTTAAA	GCTGAGGCAA	TGCTCCCTCC	CCTATCTCTC	TCTGTGTAAT
3841	TTACCATAGA	ATTAGGATGA	TTAGATTGAA	ACACATGTTG	TATGTTTTTA	AAACTACATT
3901	GCTTCATTAC	TTTCATTTTC	CGACAACATC	AACTAACAA	GAGGCAGTGT	TAAATATTTT
3961	AAATGGTGCT	ATAGCCAATG	TATTTGAATG	CTTGCACTGC	TGGTTGTGTA	TCATCAATAT
4021	GAACCTTTTA	TCCAATGACT	CAACTCTAAT	TACATCTAAG	TTAGACTTGC	TCACGTTTCAG
4081	TTTGTACAGT	TGTGTGTTGA	CTTACTATGT	TTTGAAAGTG	GTGACTTCTA	CCGAATGAGT
4141	GGAAGTTCCC	ATTGTCAAAA	AAAATAAAGA	CCTGCTTGCA	GTATTCATGT	TGACAACAGA
4201	GTAAAAGAGA	ATACTGTAAA	GAATTACTGC	AAATATTTCC	TGTTTATGTT	ATTTGCCGTT
4261	GTTTGAAGAT	ATTATAAAGG	GTTAATTGTA	TATTTATATC	ATGTGCTTTA	TCGTTTTCCC
4321	CTCATGTATC	CAAGTAATTT	TTATTTACAT	ACAACTAAAT	AAATGTTGTC	CTCTTTGAAA

4381 AAAAAAAAAA AAAAAAAAAA AAAAAAAGG GCGGCCGC

**Fbh25466pat(4419 bases) -
Nucleic Acid Db Check - All Database Hits**

Exhibit B to Accompany Declaration under
37 CFR §1.131 for US Appl. No. 10/074,547

gb U81800 HSU81800 Homo sapiens monocarboxylate transport...	384	8.4e-20	1
gb U15685 GGU15685 Gallus gallus retinal epithelial membr...	376	4.0e-19	1
gb AJ238706 DME238706 Drosophila melanogaster mRNA for ca...	292	3.9e-12	1
gb AF178954 AF178954 Mus musculus monocarboxylate transpo...	280	4.0e-11	1
gb AL009193 DMC103B4 Drosophila melanogaster cosmid clone...	292	2.4e-10	2
gb AE003423 AE003423 Drosophila melanogaster genomic scaf...	292	1.3e-09	2
gb AE003469 AE003469 Drosophila melanogaster genomic scaf...	271	3.5e-09	2
gb AR083262 AR083262 Sequence 9 from patent US 5976837.	231	4.9e-07	1

>gb|U81800|HSU81800 Homo sapiens monocarboxylate transporter (MCT3) mRNA,
complete cds.
Length = 1982

Plus Strand HSPs:

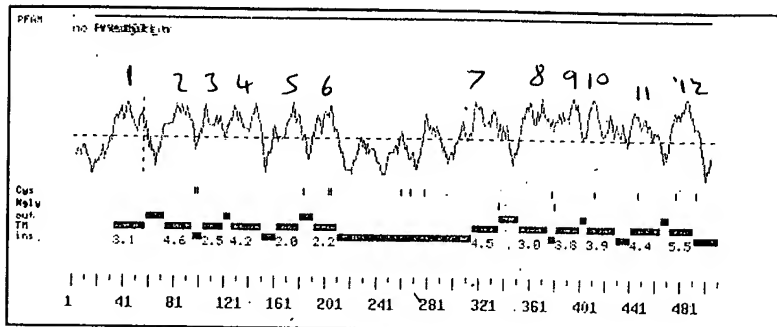
Score = 384 (108.6 bits), Expect = 8.4e-20, P = 8.4e-20

Identities = 288/552 (52%), Positives = 288/552 (52%), Strand = Plus / Plus

Query:	532	TGATGGCGGATGGGCTTGGATGATGGTGCTCTCCTCTTTCTTTGTGCACATCCTCATCAT	591
Sbjct:	110	TGACGGCGGCTGGGGCTGGGCCGTGCTCTTCGGCTGTTTCGTCATCACTGGCTTCTCCTA	169
Query:	592	GGGCTCCCAGATGGCCCTGGGTGTCCTCAACGTGGAATGGCTGGAAGAATTCCACCAGAG	651
Sbjct:	170	CGCCTTCCCCAAGGCCGTGCTGTCTTCTTCAAGGAGCTCATAAGGAGTTTGGGATCGG	229
Query:	652	CCGCGGCCTGACCGCCTGGGTGACGCTCCCTCAGCATGGGCATCACCTTGATAGTGGGCCC	711
Sbjct:	230	CTACAGCGACACAGCCTGGATCTCCTCCATCCTGCTGGCCATGCTCTACGGGACAGGTCC	289
Query:	712	TTTCATCGGCTTGTTTCATTAACACCTGTGGGTGCCGCCAGACTGCGATCATTGGAGGGCT	771
Sbjct:	290	GCTCTGCAGTGTGTGCGTGAACCGCTTTGGCTGCCGGCCCGTCATGCTTGTGGGGGGTCT	349
Query:	772	CGTCAACTCCCTGGGCTGGGTGTTGAGTGCCATGCTGCAAACGTGCATTATCTCTTCAT	831
Sbjct:	350	CTTTGCGTCGCTGGGCATGGTGGCTGCGTCTTTTGCCGGAGCATCATCCAGGTCTACCT	409
Query:	832	TACTTTTGGAGTCGCAGCTGGCCTGGGCAGCGGGATGGCCTACCTGCCAGCGGTGGTCAT	891
Sbjct:	410	CACCACTGGGGTCATCACGGGGTTGGGTTTGGCACTCAACTTCCAGCCCTCGCTCATCAT	469
Query:	892	GGTGGGCAGGTATTTCCAGAAAGAGACGCGCCCTCGCCCAGGGCCTCAGCACCACGGGGAC	951
Sbjct:	470	GCTGAACCGCTACTTCAGCAAGCGGCGCCCATGGCCAACGGGCTGGCGGCAGCAGGTAG	529

[Back to original.cgi](#)

Analysis of 25466 (510 aa)



No Signal Sequence
12 Transmembrane
Domains

Cytoplasmic N- and
C-terminals

>25466
 MYTSHEDIGVDPEDGPKDKETLKPHPNIDGGHAWRVLSPPVHILMGSONALCVLAIVE
 WLEEFHQSGRLTAMVSSLSMGITLIVGPIGLPINTCCROTATIGGLVNSLQWVLSAYA
 ANVHYLPITFGVAAGLGSCHAYLPVAVMVGYPQKRRALAQGLSTTGTGPTFLMTVLLK
 YLCATYGNRHMLTQGAIVSLNLCVCGALMRPLSPGKNPNDPGEKDVRLPAHSTESVKST
 CQGGRTTEKDDGLNEETLCDLQAOECPOQAGHRKUNSCALRLKTVSWLTHRVKGFEDW
 YSGYFGTASLPTNRNPFVAFIPWALPAYSSFPVIFHLPEIVNLYNLSEQNDVFPFLTSIIA
 IVHIFKVILOVIAIDLPCISVWNVFLANFTLVLSIFILPLMHTYAGLAVICALIGFSSG
 YFSLMPVTTEDLVGTIELANAYCIITCAHGISALLGPPFAGWITDTQKYDFSPYICGLL
 YNIGILFLLIQCIRIIEQSRKRYMDGAMV

PSORT Prediction of Protein Localization

MITDISC: discrimination of mitochondrial targeting seq
 R content: 0 Hyd Moment(75): 6.27
 Hyd Moment(95): 6.41 G content: 0
 D/E content: 2 S/T content: 2
 Score: -6.39

Gavel: prediction of cleavage sites for mitochondrial preseq
 cleavage site motif not found

NUCDISC: discrimination of nuclear localization signals
 pat4: none
 pat7: PROKTL (5) at 16
 bipartite: none
 content of basic residues: 6.5%
 NLS Score: -0.04

Final Results (k = 9/23):

66.7 %: endoplasmic reticulum
 22.2 %: mitochondrial
 11.1 %: nuclear

prediction for 25466 is end (k=9)

Start	End	Feature	Seq
387	408	Leucine zipper pattern (PS00029)	LANFTLVLSI...LPLMHTYAGL

Signal Peptide Predictions for 25466

Method	Predict	Score	Mat@
SignalP (eukaryote)	MAYBE		56

Note: amino-terminal 70aa used for signal peptide prediction

Transmembrane Segments Predicted by MEMSAT

Start	End	Orient	Score	
33	57	ins->out	3.1	TM1
73	94	out->ins	4.6	TM2
103	119	ins->out	2.5	TM3
126	149	out->ins	4.2	TM4
162	179	ins->out	2.0	TM5
191	209	out->ins	2.2	TM6
315	335	ins->out	4.5	TM7
352	374	out->ins	3.0	TM8
381	399	ins->out	3.8	TM9
406	428	out->ins	3.9	TM10
441	464	ins->out	4.4	TM11
472	490	out->ins	5.5	TM12

>25466
 HTSHEDICVDFEDGPKDKKTLKPHNIDGGWAWQNLSEFFVHILINGSQNALQVNLVE
 WLEFHQSRGLTAVVSSLSHGITLVGPFIQGLFINTCCQROTATIGGLVNSLQVLSATA
 ANVHYLFITFGVAGLGSQGNAYLPAYVGVGRYFQKRRLAQGLSTGTGFTFLN
 YLCARYGWRNAMLIGAVSLNLCCGALMRPLSPGKNPNDPGEKDVRLPAHSTESVKEST
 GQQGRTEKDGGLGNEETLCLDQAQCECPDQAGHRKNCALRIKTVSWLTHRVKGFEDW
 YSGYFGTASLPTNRNFVAFIPALFAYSSVFIPIHLPEIVNLNLSQNDVFPPLTSILA
 IVHIFGVILGVIADLPICISVWNVFLANFTLVLSIFILPLMHTYAGLAVICALQFSSG
 YFSELPVTEDLVGIHLANAYGIIICANGISALGPPFAGHIYDITQKYDFSFYICGLL
 YNIGILFLLIQPCIRIIEQSRKRYMDGAHV

Transmembrane segments for presumed mature peptide

Start	End	Orient	Score
18	39	out->ins	4.6
48	64	ins->out	2.5
71	94	out->ins	4.2
107	124	ins->out	2.0
136	154	out->ins	2.2
260	280	ins->out	4.5
297	319	out->ins	3.0
326	344	ins->out	3.8
351	373	out->ins	3.9
386	409	ins->out	4.4
417	435	out->ins	5.5

>25466 Mature
 VLVNHLSEFHQSRGLTAVVSSLSHGITLVGPFIQGLFINTCCQROTATIGGLVNSLQV
 LSAYANVHYLFITFGVAGLGSQGNAYLPAYVGVGRYFQKRRLAQGLSTGTGFTFLN
 TVLLKYLCAERYGWRNAMLIGAVSLNLCCGALMRPLSPGKNPNDPGEKDVRLPAHSTE
 SVKSTGQQGRTEKDGGLGNEETLCLDQAQCECPDQAGHRKNCALRIKTVSWLTHRVK
 GFEDWYSGYFGTASLPTNRNFVAFIPALFAYSSVFIPIHLPEIVNLNLSQNDVFPPL
 TSIATVHIFGVILGVIADLPICISVWNVFLANFTLVLSIFILPLMHTYAGLAVICALI
 GFSSGYFSLMPVTEDLVGIHLANAYGIIICANGISALGPPFAGHIYDITQKYDFSFY
 ICGLLYMIGILFLLIQPCIRIIEQSRKRYMDGAHV

Prosite Pattern Matches for 25466

Prosite version: Release 12.2 of February 1995

>PS00001|PDOC00001|ASN_GLYCOSYLATION N-glycosylation site.

Query: 345 NLSB 348

Query: 389 NPTL 392

>PS00005|PDOC00005|PKC_PHOSPHO_SITE Protein kinase C phosphorylation site.

Query: 21 TLK 23

Query: 236 SVK 238

Query: 290 TMR 292

Query: 312 TMR 314

Query: 467 TQK 469

Query: 500 SRR 502

>P500006|PDOC00006|CK2_PHOSPHO_SITE Casein kinase II phosphorylation site.

Query: 3 TSHE 6
Query: 258 TLCD 261

>P500007|PDOC00007|TYR_PHOSPHO_SITE Tyrosine kinase phosphorylation site.

Query: 294 RKGPEWY 301

>P500008|PDOC00008|MYRISTYL_N-myristoylation site.

Query: 49 GSQMAL 54
Query: 91 GLPINT 96
Query: 98 GCRQTA 103
Query: 106 GGLVNS 111
Query: 131 GVAAGL 136
Query: 137 GSGHAY 142
Query: 162 GLSTTG 167
Query: 251 GGLGNE 256
Query: 443 GIIICA 448

>P500022|PDOC00022|LEUCINE_ZIPPER Leucine zipper pattern.

Query: 387 LANPTLVLSIFILPMHTYAGL 408

>P500141|PDOC00299|MICROBODIES_CTER Microbodies C-terminal targeting signal.

Query: 508 AHV 510

Protein Family / Domain Matches, HMMer version 2

Searching for complete domains in PFAM
hmmplfam - search a single seq against HMM database
HMMER 2.1.1 (Dec 1998)
Copyright (C) 1992-1998 Washington University School of Medicine
HMMER is freely distributed under the GNU General Public License (GPL).

HMM file: /prod/ddm/seqanal/PFAM/pfam5.5/Pfam
Sequence file: /prod/ddm/wspace/orfanel/oa-script.8673.seq

Query: 25466

Model	Description	Score	E-value	N
ECT	Monocarboxylate transporter	184.0	2.5e-51	1
SUGAR_tr	Sugar (and other) transporter	-142.8	0.086	1
Presenilin	Presenilin	-256.3	4.8	1

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
Presenilin	1/1	17	336	..	1 475 []	-256.3	4.8
ECT	1/1	40	477	..	1 611 []	184.0	2.5e-51
sugar_tr	1/1	32	504	..	1 487 []	-142.8	0.086

Alignments of top-scoring domains:

Presenilin: domain 1 of 1, from 17 to 336: score -256.3, E = 4.8
 *--eeelklkY....gakhvimlfvPValCMllVvatike.vsfyeqnd
 + +lk + + + g+ m+ + + 1 + + + + + e
 25466 17 KDKKTLKPhpnidGGWAMMMVLSPPFVHILINGSQALGLVNVENLB 63
 gklilYtpPhadtptvgqkllldslInclilIsVlvmtflllvlykyrFY
 + Ph + + + + + 1 + 1 + 1 + + f + + + + +
 25466 64 E-----FHQ-----SRGLTAWVSLSLHGITLIVGPPICLPINTCGCR 100
 K..iIhgWLiLSSlIlLflfaviyLqevfkaynDQDDNNQvp.....md
 + iI.g + + + S1+ + + + + f + + + + + v + + + + m
 25466 101 QtaIIQG--LVNSLGWLSAYAAVHYLFITPG-----VAaglgagMA 141
 yptvllliwnFGvGmnciHWkppLrLqQfVLItisa....LmAlvfiky
 y+ + + + + + + + + L Q+ t + + + Lm + + ky
 25466 142 YLPAVVHVGRY-----PQKRRLAQGLSTCTGTGfgtfLNTVL-LKY 181
 LPeWteWfvLaeISvWDLVAVLcPKGpLrmlVETAgRnepiFPaliYSS
 L W mL A n + + Al+
 25466 182 LCAEYQWRN-----AMLIQAVSLNLVCGLAKRPL 212
 tyvvlitvganqeetnegtpaterrrtakstIRYyTkReaQODEFYQKIRQR
 + + + + e g p + + + kat
 25466 213 SPQKNPND-PGEKDVRLPAHSTESVKST----- 240
 raaanPdsaptshstIVEAFPSpiekkekasteesdeadsseteGmPL
 + + + + + d+ + + + L

<http://tango.mpl.com/seqanal/orfanel/25466.8670.html>

```

25466 241 -----GQ--QGRTEKDGGLGNEETLCDL 262
      vsgaSNLssdssttsqsdslaphr.teeiareqpevqsnslpsNekr
      q+ a hr+ r ++w+ l+
25466 261 QA-----QECPDQACHRKMKALRLKTVS--WLTM----- 291
      eeTAaDalhPdGeeErqwkLGLGDFiFYSvLvCkAsatasGdwmTtiACP
      v+ G+ D YS C As t F
25466 292 -----RVKGFED--WYSGYFCTASLP-----TNRMF 316
      ValLiGLclTLlllavfkALPALPisitfGIIFYPstdlvtPPmdc-+
      Va i+ +l+ Y + +v PP
25466 317 VAP-----IFWALPAY---SSFVIPPFIH 336

MCT: domain 1 of 1, from 40 to 477: score 184.0, E = 2.5e-51
      +-+sflingftdgfpksfgliffrelqeeqasnsdswsisSillavl
      sft+++ +g + + g ++ +e+eef+ s++ t+w +s+ ++ +
25466 40 SPPVHILINGSQHALG-VLNVENLEEFHQSRGLTAW--VSSLSNGIT 83
      lfaGPlasilvnrIGcrlvtiaGgllassgmvasPacniselylrfcvi
      l+GP++ + a GcR++ iGglv ssg vl+ s n+ l +tGv
25466 84 LIVGPFICLFINTCCQRTAIIGGLVNSLQWVLSAYAAHVHLYFITGVA 133
      tGIGafaiylPaiviitsYFekkRSLatgiavaGsgvgtfvlaplnpdqf
      GIG ++ ylp+ +v+++ YFek+R+LA g+++G+g+gtf+ ++l+ ++
25466 134 AGLGSGMAYLPVVMVGRYFQKRRALQGLSTTGTGPGTFLHTVLL--KY 181
      lienyGskWrgallffgmgvyvialwsvaivlncciaqalfrlpsekvk
      l + yG Mc+ +l+ g+ + l+ +c+ +gal+ rpl + +k
25466 182 LCAEYC--WRNAMLQGA-----VSLNLCVCCALMRPLSPGK-- 216
      qtklakaeepkealkskeneaseSidsirsaakaivspetpalspelt
      +P
25466 217 -----NP----- 218
      pkkdqlqkllktsctssngaklldfsvldargfllyassqslasltq
      +++ + + l +s++s+ s g+q
25466 219 -----NDPGEKDVRLGLPAHSTESVKSTGQQ 243
      lfipgsiflvnfekslgenlssvkskeafllsilgdsadkeggfifir
      + ++ k+ g + e+ l+ d+ g + + + +
25466 244 GRTE-----SKDGG-----LGNEETLCDLQAG-BCPDQACHRKMK 279
      patllsflgfvanketksncrpvllyllslcniavvngilrslasala
      + + + + + + + + + + + + + + + +
25466 280 LRI-----LKTVM--LTMVRKGF-----DWYSGYFG 306
      gsrkekksaidkielkstfvgllfslfifgvgfsgkavvilalagfll
      + + l+ + + + + + + + + + + + + + + +
25466 307 T-----ASLPTNRMFVAFIPW-----ALPAYSSPV 331
      fsilyaipvvgllqkyssalGltetdastlleaivlniiGrplaGllaDk
      + + + + + v l + + s+ ++ l+ + + + + i G + + + G + aD
25466 332 IPIHLPEIVNLNLSQNDVF-----PLTSITAIHVIFGKVLGVIAIDL 376
      ckrklaiynslilcClfvafPlatfifglafyvcvlfGsvifllayaf
      + + + + + + + + + + + + + + + + + + + + + +
25466 377 PCISVMVFLANFTLVLSIFILPLMGTIYAGLAVICALI----- 415
      kgfckgsyialtsviavdltdgldksnafgllllfqGvatlvGppiagll
      gf+ g y+ + + + + + + + + + + + + + + + + + + + + +
25466 416 -GFSSG-YFSLMPVVTEDLVGIEHLANAYGIIICANGISALGGPPFACMI 463
      kdltsykvafyfae-+
      +d+ + y+ sfy+ +
25466 464 YDITQKYDFSPYIC 477

sugar_tr: domain 1 of 1, from 32 to 504: score -142.8, E = 0.086
      +-+valvaalGqfifGyDtgvggglalidifrfglitssgalaeqy
      + + + + + l+ + + + + + + + + + + + + + + + + +
25466 32 MAMWNVLS-----PFVHILINGSQHALGVLENLEEFHQSRGL--- 71
      stvlctglvsvfflGrllGslfaGklgdrfGRkklialvlvfiGalls
      t++ + s + li + E G + + + G + + + + + + + + + + + +
25466 72 ----TAMVSSLSNGITLIVGPFICLFINTCCQRTAIIGGLVNSLQWVLS 117
      gaapgytTiGlvafyllivGRvlgGlgvGgasvlpvafisEIAFkaIRGa
      + + + + + yvli + v +Glg G a+ + + + + + + + + + + +
25466 118 AYAAV-----HYLFIPTGVAAGLGSCKAYLPVVMVGRY--FQKRR 150
      lgslyqlaitiGllvAaiglglnktndnsalnsqWRipglqlvpall
      l+ + l+ + + + + + + + + + + + + + + + + + + + + +
25466 159 LA--QGLSTTGTGPTFLMTVLLKYLCA-----EYQWRNAMLQGAVALN 201
      lligllfIP.....ESPRvLveghkleeAre
      l+ + + + + + + + + + + + + + + + + + + + + + + + + + + +
25466 202 LCVCCALMRPLspgknpndpgekdvrglpahstESVKSTGQQGRTEKDG 251
      vLak...lrgv.....edvdqelqeikaeeagveee.k
      L + + l+ + + + + + + + + + + + + + + + + + + + + +
25466 252 GLGNeetLCDLqaeqcdqaghrkmcALRLKTVSWLTMVRKGFEDWY 301
      agkasvgelfggrtrpkvrqclmgvmlqafqGltGmaifYsptifks
      + + + + + + + + + + + + + + + + + + + + + + + + + + + +
25466 302 SGYPGTASLPTN-----RMFVAFIPWALPAYSSFPVIFPIHLPEIVNL 343
      .....vGvdsraallvtliugvvNfvfTlvaljflvDrIGRRpili
      + + + + + + + + + + + + + + + + + + + + + + + + + + + +
25466 344 ynlsqndvFPPLTS-----ITAIHVIFGKVLGVIAIDLPCISVM--NVFL 386
      lGaagmleclilgasigvalilinkpkdpstaagivalvtilifiaFP
      + + + + + + + + + + + + + + + + + + + + + + + + + + + +
25466 387 LANFTLVLSIFIL-----PLMGTIY-----AGLAVICALIG---F 417
      algwGpwpvileElFPkvrskalataaanlanfligflfpyitgai
      g+ + + + + + + + + + + + + + + + + + + + + + + + + + + +

```

```
25466 418 SSGYFSLMPVVTEDLVGIEHLNAYGIIICANGISALLGPPFAGMIYDIT 467
      glalggvvlvflapllvflilvfflvPETKgrtLeeieelf<+
      y f ++ + +l + ++f+++ +r +e+ + +
25466 468 Q----KYDFSPYL-CGLLYMIGILFLLIQPCIRIIEQSRKY 504

//
Searching for complete domains in SMART
hmmpfam - search a single seq against HMM database
HMMER 2.1.1 (Dec 1998)
Copyright (C) 1992-1998 Washington University School of Medicine
HMMER is freely distributed under the GNU General Public License (GPL).
-----
HMM file: /ddm/robison/smart/smart/smart.all.hmms
Sequence file: /prod/ddm/wspace/orfanel/oa-script.8673.seq
-----
Query: 25466







Scores for sequence family classification (score includes all domains):
Model Description Score E-value N
-----
[no hits above thresholds]

Parsed for domains:
Model Domain seq-f seq-t hm-f hm-t score E-value
-----
[no hits above thresholds]

Alignments of top-scoring domains:
[no hits above thresholds]

//
```

ProDom Matches

ProdomId	Start	End	Description	Score
View Prodom 38068  	345	461	p99.2 (2) O06142(1) O30820(1) // PROTEIN ADENYL CYCLASE	73
View Prodom 102799  	393	467	p99.2 (1) YU45_MYCTU // HYPOTHETICAL 41.9 KD PROTEIN CY339.45C TRANSMEMBRANE	80
View Prodom 84838  	440	494	p99.2 (1) Q23113_CAEEL // SIMILAR TO SYNAPTIC VESICLE AMINE TRANSPORTER	74
ProdomId	Start	End	Description	Score

View Prodom 84838

```
>84838 p99.2 (1) Q23113_CAEEL // SIMILAR TO SYNAPTIC VESICLE AMINE TRANSPORTER
Length = 85

Score = 74 (31.1 bits), Expect = 0.053, P = 0.052
Identities = 18/57 (31%), Positives = 31/57 (54%)

Query: 440 NAYGIIICANGISALLGPPFAGMIYDITQKYDFSPYICGLLYMI--GILFLLIQPCI 494
      +* CI + + L+GPP+ G +Y + K + F + LL + I F+++QP I
Sbjct: 14 SAMGIALGGALGVLVGPYPYGGLLYQMSGR-ELPPVLLALLALFDGSIQPMVLQPKI 69
```

View Prodom 102799

```
>102799 p99.2 (1) YU45_MYCTU // HYPOTHETICAL 41.9 KD PROTEIN CY339.45C
TRANSMEMBRANE
Length = 155

Score = 80 (33.2 bits), Expect = 0.55, P = 0.43
Identities = 25/78 (32%), Positives = 37/78 (47%)

Query: 393 VLSIFILPLMHT-YAGLAVICALI-GPSSGYF-SLMPVVTEDLVGIEHLNAYGIIICAN 449
      +LSI ++ T + VI L+ C + G E + V DLV A+AYG+
Sbjct: 48 ILSILVLFAPTDNVTVVIGTLVWGAAVGICQESTLRGVVADLVASPRASAYGVFAAGL 107

Query: 450 GISALLGPPFAGMIYDIT 467
      C + G GW+YDI+
Sbjct: 108 GAATAGGGALIGWLVDIS 125
```

View Prodom 38068

```
>38068 p99.2 (2) O06142(1) O30820(1) // PROTEIN ADENYL CYCLASE
Length = 218

Score = 73 (30.8 bits), Expect = 8.1, P = 1.0
```

<http://range.mpi.com/seqans/orfanel/25466.8670.html>

Identities = 41/135 (30%), Positives = 62/135 (45%)

Query: 345 NLSEQNDVPLTSTIIAIVHI--FGKVLGVADLPCISVWNVFLANFTLV--LSIF--- 397
+++ + V +T+ +A+V F + L A I++ NVF F +V L F
Sbjct: 15 SVARRQRLTITANLAVVVTGSAFMQLATGAGGWYIALINVPTAVTFATVPLLRPGGL 74

Query: 398 ILPLMHTYAGLAVICAL-ICPSSG-----YPSLMPVVTEDLVGIEHLNAYGIIICAN 449
+ PL T+ G A + IG+ G +P + + LVGIEH A A G+ A
Sbjct: 75 VAPL--TPIGTAVVAIPAIGWDVGTGAGAQPFLLVAAALVLLVGIENTALAVGLAAVAA 132

Query: 450 G-ISAL--LGPPFAG 461
G + AL L PP G
Sbjct: 133 GLVIALEFLVPPDTG 147

CLUSTAL W (1.74) multiple sequence alignment

Exhibit D to Accompany Declaration under
37 CFR §1.131 for US Appln. No. 10/074,547

```

Fbh25466pat    MYTSHEDIGYDFEDGPKDKKTLKPHPNIDGGWAWMMVLSSFFVHILIMGSQMALGVLNVE
O15374          -----MLKRE-GKVQPYT-----KTLDGGWGWMIVIHFFLVNVFVMGMTKTFAIFFVV
                  * * * : * :*****:*:* :*:***** :*: *

Fbh25466pat    WLEEFHQSRGLTAWVSSLSMGITLIVGPFIFLINTCGCRQTAIIGGLVNSLGWVLSAYA
O15374          FQEEFEGTSEQIGWIGSIMSSLRFCAGPLVAIICDILGEKTTTILGAFVVTGGYLISSWA
                  : **** : :*:*:* : : :*:*:*:* : * :*:*:*:* :*:*:*:*

Fbh25466pat    ANVHYLFITFGVXAGLGSGMAYLPAVXMGVGRYFQKRRALAQGLSTTGTGFGTFLMTVLLK
O15374          TSIPFLCVTMGLLPGLGSAFLYQVAAVVTTKYFKRLALSTAIARSGMGL-TFLLAPFTK
                  : : * :*:* :*:*:* : * * : :*:* * * : : * * :*:* : *

Fbh25466pat    YLCAEYGWRNAMLIQGAVSLNLCVCGALMRPLSPGKNPNDPGEKDVRG-LPAHSTESVKS
O15374          FLIDLWDWTGALILFGAIALNLVPSSMLLRPI-HIKSENNSGIKDKGSSLSAHGPEAHAT
                  : * *.* :*:* :*:*:* : * :*:* : * :*.* * * :*:*:* :

Fbh25466pat    TGOQGRTEE---KDG---GLGNE-ETLCDLQAQECPDQAGHRKNMCALRILKTVSWLTMR
O15374          ETHCHETEESTIKDSTTQKAGLPSKNLTVSQNQSEEFYNGPNRN---RLLKSDDEESDK
                  : :*** ** : * :* * * :* :* * :* :* : :

Fbh25466pat    VRKGFEDWYSGYFGTASLFTNRMFVAFIFWA-LFAYSSFVIPFIHLPEIVNLYNLSEQND
O15374          VIS---WSCKQLFDISLFRNPFFYIFT-WSFLLSQLAYFIPTPHLVARAKTLGID-IMD
                  * : * : :*** * : * * :* : :* * :* : : :*

Fbh25466pat    VFPLTSIIAIVHIFGKVLGVIALDPCISVWNVFLANFTLVLSIFILPLMHTYAGLA--
O15374          ASYLVSAGILETVSQIISGWVADQNWIKYHYHKSYLILCGITNLLAPLATTFPLLMTY
                  : *.* :*:* :*:*:* :* * :* : : : : :* * :*

Fbh25466pat    VICALIGFSSGYFSLMPVVTEDLVGIEHLANAYGIIICANGISALLGPPFAGWIYDITQK
O15374          TICFAI-FAGGYLALILPVLVDLCRNSTVNRFLGLASFFAGMAVLSGPPPIAGWLYDYTQT
                  :* * * :*:*:* : * * : : * :* :* :*:*:* * *

Fbh25466pat    YDFSFYICGLLYMIGILFLLIQPCIRIIEQSRRKYMDGAHV
O15374          YNGSFYFSGICYLLSSVSFFFPV---LAERWKNSLT-----
                  * : *** :* :*:* : : :* : * :* :*

```